Is Kriging-techniek geschikt gereedschap voor meetnetoptimalisatie?

Kriging-techniek wordt als ruimtelijke interpolatietechniek in vele vakgebieden toegepast. Deze interpolatietechniek sluit aan op situaties waarin per waarnemingslocatie 1 waarneming beschikbaar is.
Dank zij zijn succes bestaat ook de wens om de methode in andere situaties en voor andere vraagstellingen toe te passen, bijvoorbeeld bij de optimalisatie van grondwaterstijghoogtemeetnetten.
De Kriging-techniek wordt voor deze afwijkende context op zijn merites beoordeeld. Geconcludeerd kan worden dat de resultaten van de huidige meetnet-optimalisaties met behulp van Kriging-techniek niet veel meer zijn dan gecompliceerde vertalingen van ongeldige vooronderstellingen.



DR. IR. J. W. DE KWAADSTENIET
RIVM, Laboratorium voor Bodem- en Grondwateronderzoek

1. Inleiding

In de jaren zestig werd in de Ecole des Mines te Fontainebleau de wiskundige basis gelegd voor de ruimtelijke interpolatietechniek die bekend staat als de Kriging-techniek, genoemd naar de in de Zuidafrikaanse mijnbouw werkzame initiator D. G. Krige. Deze techniek is voortgekomen uit de wens om meer gericht onderzoek te kunnen uitvoeren naar nieuwe vindplaatsen voor goud, maar heeft inmiddels in een veel ruimere context binnen de mijnbouw en in belangrijke mate ook daarbuiten zijn waarde bewezen.
Wiskundig gezien wordt in de Kriging-techniek gebruik gemaakt van kansrekening, waarmee vanuit de waarnemingen op een aantal bemeten locaties op elke niet-bemeten locatie een schatting van de betreffende variabele kan worden gegeven. De gebruikte kansrekening resulteert daarbij in betrouwbaarheidsinformatie met betrekking tot die schatting in de vorm van een standaarddeviatie. De wiskundige eenvoud waarmee de schattingen en de betrouwbaarheidsinformatie kunnen worden verkregen, verklaart in belangrijke mate de ruime toepassing die de methode heeft gekregen.
Een hoeksteen van de methode wordt gevormd door het semi-variogram (of in een algemenere formulering de gegeneraliseerde covariantie) waarmee de variantie- en correlatiestructuur van de variabele onder studie wordt beschreven. Hiermee wordt vervolgens de betrouwbaarheidsinformatie van de schattingen berekend. Daarbij wordt er in de meest simpele situatie (1 semi-variogram) vanuit gegaan dat (a) de variantie overal in het beschouwde gebied dezelfde waarde heeft en (b) de ruimtelijke correlatie in de betreffende variabele behorend bij een tweetal locaties slechts een functie is van de afstand van de betreffende locaties.
(In meer gecompliceerde toepassingen wordt met semi-variogrammen of gegeneraliseerde covarianties per deelgebied of per windrichting gewerkt). De bij de schattingen voor de onbemeten locaties berekende betrouwbaarheidsinformatie zal nu sterk afhangen van de gebruikte semi-variogram- resp. gegeneraliseerde covariantiefuncties, terwijl ongeldigheid van de daartoe benodigde vooronderstellingen zal resulteren in ongeldigheid van deze betrouwbaarheidsinformatie. In dit verband is het van belang er op te wijzen dat de beschreven hoeksteen van de Kriging-techniek ook meteen zijn Achillespees vormt en wel om de volgende redenen.
1. De gemaakte ingrijpende vooronderstellingen zijn vaak niet of nauwelijks op hun geldigheid te toetsen (samenhangend met het feit dat in vele toepassingsgebieden per definitie slechts 1 gerealiseerde waarneming per waarnemingslocatie beschikbaar is).
2. In het geval dat aan de vooronderstellingen zou zijn voldaan, is vervolgens geen statistische maat aanwezig voor het vastleggen van een functieverband voor een semi-variogram of gegeneraliseerde covariantie.
3. In het geval dat de waarnemingen zouden resulteren in eenduidige functieverbanden voor semi-variogrammen of gegeneraliseerde covarianties, zou in het berekenen van betrouwbaarheidsmaten voor de schattingen met de Kriging-techniek de waargenomen variatie rond deze functieverbanden ten onrechte niet worden verwerkt.
De genoemde ernstige tekortkomingen aan de basis van de Kriging-techniek zijn terug te voeren tot het feit dat weliswaar sprake is van een probabilistisch model (dat wil zeggen van een model waarin met waarschijnlijkheidsrekening wordt gewerkt) maar niet van een statistisch model, dat wil zeggen niet van een model dat er op gericht is de betrouwbaarheidsinformatie voor de schattingen zo efficiënt mogelijk te ontlenen aan de basale variatie die in het waarnemingsmateriaal aanwezig is.
De bovenstaande kritiek op de Kriging-techniek laat echter onverlet dat toepassing ervan veelal een belangrijke stap vooruit betekent ten opzichte van hetgeen tot dan toe gebruikelijk is, gezien het feit dat een schatting steeds wordt vergezeld van een eerste indicatie over de betrouwbaarheid ervan. Daarbij dient echter wel expliciet te worden aangetekend dat (1) Kriging-techniek is ontwikkeld als een ruimtelijke *interpolatie*techniek, die (2) aansluit op situaties waarin men per waarnemingslocatie slechts over *1 waarneming* aan de variabele beschikt en dat de techniek opnieuw op zijn merites en relevantie moet worden beoordeeld indien gebruik in een andere dan de in (1) en (2) aangegeven context wordt overwogen. Een voorbeeld van een dergelijke afwijkende context, dat zijn belang ontleent aan de activiteiten die dienaangaande in provinciaal verband plaatsvinden, is de optimalisatie van meetnetten voor grondwaterstijghoogte. Waardebepaling van de Kriging-techniek in deze context is het thema in het navolgende.

2. Uitwerking

De basisgedachte die aan meetnet-optimalisatie ten grondslag ligt, laat zich als volgt omschrijven.
Welke meetpunten van een gegeven verzameling van meetpunten binnen een bepaald gebied moeten worden gehandhaafd en waar moeten eventueel nieuwe meetpunten worden bijgeplaatst, wil men binnen dit gebied en met een bepaalde betrouwbaarheid een compleet ruimtelijk beeld van de desbetreffende variabele (in dit geval de stijghoogte) kunnen geven op elk van de waarnemingstijdstippen?
De in de Kriging-techniek gebruikelijke handelwijze om vanuit de meetwaarden tot een compleet ruimtelijk beeld van een variabele met betrouwbaarheidsinformatie te komen, is een iteratieve, waarbij men in elke iteratiestap de te beschrijven variabele gesplitst denkt in een deterministisch deel (de ruimtelijke trend of drift) en een ruisdeel, welke laatste wordt beschreven met één of meerdere semi-variogrammen of gegeneraliseerde covarianties. In de procedure wordt gebruik gemaakt van 1 waarneming per waarnemingslocatie. Nu is echter, onafhankelijk van de wijze waarop de iteratieprocedure wordt ingevuld, daarbij altijd de fundamentele kanttekening te plaatsen dat op basis van het beschouwde waarnemingsmateriaal (1 waarneming per waarnemingslocatie) er geen keuze mogelijk is tussen het deel van de beschouwde ruimtelijke variabele dat deterministisch beschreven moet worden en het deel dat moet worden beschreven met kansuitspraken (het ruisdeel). Pas bij specificaties ten aanzien van één van de beide delen zal men specificaties kunnen geven van het andere deel. Wordt het deterministische deel de mogelijkheid gelaten gedetailleerd het ruimtelijke gedrag te volgen, dan zal de bijdrage van het ruisdeel relatief gering zijn, terwijl bij een grof deterministisch deel een grotere component van het ruimtelijk gedrag in het ruisdeel wordt ondergebracht. De verdeling tussen deterministisch deel en ruisdeel die men kiest, zal uiteindelijk doorwerken tot in de geschatte waarden en de berekende betrouwbaarheden, zodat eenduidigheid

dienaangaande niet wordt verkregen.
De grote mate van subjectiviteit in de gekozen verdeling resulteert bovendien in een maar zeer beperkte waarde die kan worden toegekend aan de twee delen afzonderlijk (dat wil zeggen aan deel 1: de uiteindelijk vastgelegde deterministische beschrijving en deel 2: de uiteindelijk vastgelegde beschrijving van semi-variogrammen of gegeneraliseerde covarianties). Zo is het bijvoorbeeld weinig passend om te besluiten tot het opheffen van meetpunten die een afstand hebben tot een ander meetpunt die kleiner is dan de autocorrelatiestraal volgens een semi-variogram dat berekend is op de bovengeschetste wijze en meetpunten toe te voegen waar die afstand groter is dan die autocorrelatiestraal.

Nu is echter in het geval van meetnet-optimalisatie de situatie fundamenteel anders dan de situatie waarin Kriging-techniek het eerst werd toegepast. Had men in de mijnbouw per waarnemingslocatie slechts één waarneming, nu beschikt men per waarnemingslocatie over een *tijdreeks* aan waarnemingen en dus over een extra bron van informatie die in de meetnetoptimalisatie kan worden betrokken. Als een eerste stap in die richting kan uit de tijdreeksen informatie worden ontleend over de ruimtelijke variantie- en correlatiestructuur in het over een tijdsperiode gemiddelde beeld van de variabele, waarbij elke binnen die tijdsperiode gerealiseerde waarneming zijn gecorreleerde en kwantificeerbare bijdrage levert tot dit beeld. Bij het RIVM werd op deze wijze van het langjarig gemiddelde stijghoogtebeeld in het oosten van de provincie Utrecht de ruimtelijke variantie- en correlatiestructuur bepaald met als meest wezenlijke berekende eigenschappen:
(a) ruimtelijke inhomogeniteit in variantie over afstanden van enkele kilometers en
(b) zeer grote autocorrelaties over het gebied (autocorrelaties van 0.8 over afstanden van 20 km). Nu is de variantie- en correlatiestructuur van het gemiddelde beeld weliswaar niet identiek aan de betreffende structuur van de variabele zelf, maar wel kan worden gesteld dat (1) de berekende structuur over het stijghoogtebeeld de eerste is die met een minimum aan vooronderstellingen wordt ontleend aan het stelsel waarnemingen en dat (2) bij overdenking van de onder (a) en (b) weergegeven resultaten blijkt dat deze goed aansluiten bij de verwachtingen ten aanzien van het gedrag van het grondwater in dergelijke gebieden.

Op grond van (1) en (2) doet men er dan ook goed aan bij gedachtenvorming over de grondwaterstijghoogte te denken in termen van inhomogniteit in variantie en grote autocorrelaties.

De beantwoording van de eerder genoemde vraag naar specificatie van de splitsing van het ruimtelijk gedrag van de grondwaterstijghoogte in een deterministisch deel en een ruisdeel kan nu nader vorm worden gegeven via concretere ideeën over het ruisdeel, gebruik makend van de informatie die aanwezig is in de tijdreeksen op de bovengeschetste wijze. Daarbij blijkt nu (o.a. gelet op inhomogeniteit in variantie) dat het niet voor de hand ligt om de beschrijving van het ruisdeel te vangen in termen van semi-variogrammen of gegeneraliseerde covarianties. Maar verrassender nog zijn de gevolgen van de grote autocorrelaties. Ruwweg zal het op voorhand duidelijk zijn dat er iets niet klopt als we in extremum stellen dat 1 meetpunt in een straal van 50 km voldoende is voor het beschrijven van het ruimtelijke gedrag van het grondwaterstijghoogtebeeld omdat binnen die straal nog een aanzienlijke autocorrelatie in het verschijnsel aanwezig is.
Indien we (nader geformuleerd) van de ruimtelijke variabele de volgende drie onafhankelijke aspecten in de beschouwingen betrekken:
aspect 1. = het deterministische deel van de variabele (= de ruimtelijke trend of drift = het deel dat men gewoonlijk modelleert met behulp van mathematisch fysische modellen = (meer statistisch geformuleerd) het ruimtelijke verloop van de verwachtingswaarde van de met behulp van kansrekening beschreven variabele),
aspect 2. = het ruimtelijke verloop van de variatie rond de verwachtingswaarde (= het ruimtelijke gedrag van de variantie);
aspect 3. = de ruimtelijke samenhang in de variatie rond de verwachtingswaarde (= de ruimtelijke correlatie in de variabele), dan kunnen we het volgende stellen.
Aspect 3 is overal eenvoudig voorspelbaar, met als gevolg dat de informatie die elk meetpunt van een meetnet afzonderlijk geeft hoofdzakelijk bestaat uit informatie omtrent de verwachtingswaarde van de variabele in dat punt (aspect 1) en de variatie rond de verwachtingswaarde in dat punt (aspect 2). Dit betekent dat een meetnetoptimalisatie op basis van aspect 3, zoals bijvoorbeld in het bovengeschetste simpele voorbeeld werd aangegeven maar ook in geavanceerdere Kriging-toepassingen een belangrijke rol speelt, niet erg passend is aangezien elk meetpunt juist informatie geeft over de van 3 onafhankelijke aspecten 1 en 2.

Tot besluit is het van belang om na te gaan hoe de terecht gevoelde behoefte om de gangbare deterministische benadering van de grondwaterstroming van betrouwbaarheidsinformatie te voorzien uiteindelijk uitpakt ten aanzien van de ruimtelijke beschrijving van de grondwaterstijghoogte. De grote mate van autocorrelatie blijkt dan zeer hoge eisen te gaan stellen aan de ruimtelijke beschrijving van het deterministische deel (aspect 1), in die zin dat een beschrijving van het deterministische deel pas als passend kan gelden (en anders als niet-passend wordt verworpen) indien (ruwweg) de afwijkingen van de waarnemingen in de meetpunten ten opzichte van deze beschrijving (a) allemaal binnen de toegestane variatie per meetpunt blijven (aspect 2) en (b) allemaal hetzelfde teken hebben (aspect 3). (ad. (b).
De deterministische beschrijving mag dus niet zo grof zijn dat zowel positieve als negatieve residuen worden gevonden.)
Vanuit de statistiek wordt dus het belang van de vanuit het verleden gangbare, gedetailleerde benadering van het grondwaterstijghoogtebeeld onderstreept, zoals deze tot uitdrukking is gekomen in de oorspronkelijke opzet van de meetnetten.

Concluderend kunnen we aan de hand van het bovenstaande stellen dat de resultaten van de huidige meetnetoptimalisaties met behulp van Kriging-techniek niet veel meer zijn dan gecompliceerde vertalingen van ongeldige vooronderstellingen (zoals (1) de gekozen splitsing in een deterministisch deel en een ruisdeel, (2) de optimalisatie vooral gericht op het minst interessante aspect (de correlatie), (3) de vooronderstellingen ten aanzien van homogeniteit in variantie en correlatie als functie van afstand en richting), aan welke conclusie een eerste antwoord op de uitgangsvraagstelling kan worden ontleend.

• • •

Regio bijeenkomst Watervoorziening van Limburg en Noord-Brabant vanuit de Maas

KIVI-sectie waterbeheer en de Vereniging van Ingenieurs in Zuid-Limburg organiseren op woensdagmiddag 26 februari 1986 bij het Waterschap Zuiveringschap Limburg te Roermond een bijeenkomst over de Watervoorziening van Limburg en Noord-Brabant vanuit de Maas. Sprekers zijn:
ir. G. Bleichrodt (RWS directie Limburg): 'De Maas als bron voor de regionale watervoorziening';
ir. K. J. Provoost (PWS Noord-Brabant): 'Watervoorziening in het provinciaal Waterhuishoudingsplan';
ir. J. S. J. Dragt (Waterschap De Aa): 'Watervoorziening op waterschapsniveau';
ir. T. P. Koeman (Landinrichtingsdienst Limburg): 'Betekenis van wateraanvoer voor zandgebieden'.